

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/007,306

DATE: 06/09/98
 TIME: 08:38:51

INPUT SET: S3103.raw

**This Raw Listing contains the General
 Information Section and up to the first 5 pages.**

SEQUENCE LISTING

(1) General Information

ENTERED

(i) APPLICANT: Hillman, Jennifer L.

(ii) TITLE OF THE INVENTION: A NOVEL GUANOSINE MONOPHOSPHATE REDUCTASE

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

(B) STREET: 3174 Porter Drive

(C) CITY: Palo Alto

(D) STATE: CA

(E) COUNTRY: USA

(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette

(B) COMPUTER: IBM Compatible

(C) OPERATING SYSTEM: DOS

(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/007,306

(B) FILING DATE:

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/774,169

(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.

(B) REGISTRATION NUMBER: 36,749

(C) REFERENCE/DOCKET NUMBER: PF-0182 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415-855-0555

(B) TELEFAX: 415-845-4166

(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

RAW SEQUENCE LISTING PATENT APPLICATION US/09/007,306

DATE: 06/09/98
TIME: 08:38:53

INPUT SET: S3103.raw

```

47      (i) SEQUENCE CHARACTERISTICS:
48          (A) LENGTH: 366 amino acids
49          (B) TYPE: amino acid
50          (C) STRANDEDNESS: single
51          (D) TOPOLOGY: linear
52
53      (vii) IMMEDIATE SOURCE:
54          (A) LIBRARY: Consensus
55          (B) CLONE: Consensus
56
57      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
58
59      Met Thr Cys Cys Leu Pro Ala Leu Arg Phe Ile Ala Thr Pro Arg Leu
60      1          5          10          15
61      Ser Ala Met Pro His Ile Asp Asn Asp Val Lys Leu Asp Phe Lys Asp
62      20          25          30
63      Val Leu Leu Arg Pro Lys Arg Ser Thr Leu Lys Ser Arg Ser Glu Val
64      35          40          45
65      Asp Leu Thr Arg Ser Phe Ser Phe Arg Asn Ser Lys Gln Thr Tyr Ser
66      50          55          60
67      Gly Val Pro Ile Ile Ala Ala Asn Met Asp Thr Val Gly Thr Phe Glu
68      65          70          75          80
69      Met Ala Lys Val Leu Cys Lys Phe Ser Leu Phe Thr Ala Val His Lys
70      85          90          95
71      His Tyr Ser Leu Val Gln Trp Gln Glu Phe Ala Gly Gln Asn Pro Asp
72      100          105          110
73      Cys Leu Glu His Leu Ala Ala Ser Ser Gly Thr Gly Ser Ser Asp Phe
74      115          120          125
75      Glu Gln Leu Glu Gln Ile Leu Glu Ala Ile Pro Gln Val Lys Tyr Ile
76      130          135          140
77      Cys Leu Asp Val Ala Asn Gly Tyr Ser Glu His Phe Val Glu Phe Val
78      145          150          155          160
79      Lys Asp Val Arg Lys Arg Phe Pro Gln His Thr Ile Met Ala Gly Asn
80      165          170          175
81      Val Val Thr Gly Glu Met Val Glu Glu Leu Ile Leu Ser Gly Ala Asp
82      180          185          190
83      Ile Ile Lys Val Gly Ile Gly Pro Gly Ser Val Cys Thr Thr Arg Lys
84      195          200          205
85      Lys Thr Gly Val Gly Tyr Pro Gln Leu Ser Ala Val Met Glu Cys Ala
86      210          215          220
87      Asp Ala Ala His Gly Leu Lys Gly His Ile Ile Ser Asp Gly Gly Cys
88      225          230          235          240
89      Ser Cys Pro Gly Asp Val Ala Lys Ala Phe Gly Ala Gly Ala Asp Phe
90      245          250          255
91      Val Met Leu Gly Gly Met Leu Ala Gly His Ser Glu Ser Gly Gly Glu
92      260          265          270
93      Leu Ile Glu Arg Asp Gly Lys Lys Tyr Lys Leu Phe Tyr Gly Met Ser
94      275          280          285
95      Ser Glu Met Ala Met Lys Lys Tyr Ala Gly Gly Val Ala Glu Tyr Arg
96      290          295          300
97      Ala Ser Glu Gly Lys Thr Val Glu Val Pro Phe Lys Gly Asp Val Glu
98      305          310          315          320
99      His Thr Ile Arg Asp Ile Leu Gly Gly Ile Arg Ser Thr Cys Thr Tyr

```

RAW SEQUENCE LISTING PATENT APPLICATION US/09/007,306

DATE: 06/09/98
TIME: 08:38:55

INPUT SET: S3103.raw

```

100                               325                               330                               335
101 Val Gly Ala Ala Lys Leu Lys Glu Leu Ser Arg Arg Thr Thr Phe Ile
102                               340                               345                               350
103 Arg Val Thr Gln Gln Val Asn Pro Ile Phe Ser Glu Ala Cys
104                               355                               360                               365

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1843 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
- (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

120 CTAAATCAGC CTCTTGCCCC ATTGCTCTTT GCAGGGGTAG AAGAAGGAAG TGTAGCGGGG 60
121 TAAGGAATGC ACCGTCAGGG TCTCTCACAA CCCTTTCCCA GCTCTCCTCC CCAACAAACA 120
122 GTACCTGGGA TGGAGCCCTA GGGTAATCGC AGCCACGGGA TGGGTCGAGG TGACAGGCTT 180
123 CAGGGACCAC ACTTCGGCCT TTGCCCGACC TTCCACAAC TAAGCGAAGA GAGGCCACCA 240
124 GCCGTAACAG GGC GTTAAAG CCCAGGGGAA GATTGGTCCT TATGACTTGC TGCCTTCCAG 300
125 CCCTCAGATT CATCGCTACC CCGAGGCTAA GCGCCATGCC TCATATTGAC AACGATGTGA 360
126 AACTGGACTT CAAGGATGTC CTTTTGAGGC CCAAACGCAG TACCCTTAAG TCTCGAAGTG 420
127 AGGTGGATCT CACAAGATCC TTTTCATTTT GGAACTCAA GCAGACATAC TCTGGGGTTC 480
128 CCATCATTGC TGCCAATATG GATACTGTGG GCACCTTTGA GATGGCCAAG GTTCTCTGTA 540
129 AGTTCTCTCT CTTCACTGCT GTCCATAAGC ACTATAGCCT CGTTCAGTGG CAAGAGTTTG 600
130 CTGGCCAGAA TCCTGACTGT CTTGAGCATC TGGCTGCCAG CTCAGGCACA GGCTCTTCTG 660
131 ACTTTGAGCA GCTGGAACAG ATCCTGGAAG CTATTCCCCA GGTGAAGTAT ATATGCCTGG 720
132 ATGTGGCAAA TGGCTACTCT GAACACTTTG TTGAATTTGT AAAAGATGTA CGGAAGCGCT 780
133 TCCCCAGCA CACCATCATG GCAGGGAATG TGTAACAGG AGAGATGGTA GAAGAGCTCA 840
134 TCCTTTCTGG GGCTGACATC ATCAAAGTGG GAATTGGGCC AGGCTCTGTG TGTACTACTC 900
135 GGAAGAAAAC TGGAGTGGGG TATCCACAGC TCAGCGCAGT GATGGAGTGT GCAGATGCTG 960
136 CTCATGGCCT CAAAGGCCAC ATCATTTCAG ATGGAGGTTG CAGCTGTCCT GGGGATGTGG 1020
137 CCAAGGCTTT TGGGGCAGGA GCTGACTTCG TGATGCTGGG TGGCATGCTG GCTGGGCACA 1080
138 GTGAGTCAGG TGGTGAGCTC ATCGAGAGGG ATGGCAAGAA GTACAAGCTC TTCTATGGAA 1140
139 TGAGTTCTGA AATGGCCATG AAGAAGTATG CTGGGGGCGT GGCTGAGTAC AGAGCCTCAG 1200
140 AGGGAAAGAC AGTGGAAGTT CTTTTTAAAG GAGATGTGGA ACATACCATC CGAGACATCC 1260
141 TAGGAGGGAT CCGCTCTACG TGTACCTATG TGGGAGCAGC TAAGCTCAA GAGTTGAGCA 1320
142 GGAGAACTAC CTTTCATCCG GTCACCCAGC AGGTGAATCC AATCTTCAGT GAGGCGTGCT 1380
143 AGACCTGAGC AGTTCTACCC TCCCAAGGCA CCAGTACTCT ACCATGGGGC ATCCCAAGTG 1440
144 GGGTCCTCAC CCATCCCAGC TACTGCAGCT CTGTATTACT TTGTCAATTC CTGTTGTCTC 1500
145 ACTCCTGAGG GCTCCTGCAG TAACTCTGTA CTTCTCTATC TGCACACACA AAATGCCCAA 1560
146 GGCACCTACT GGGGAGGAAG CAAGGAAGCA AACAGTCTGA GGAAATGATG CAAGAAAATC 1620
147 AAATGGGAAT CTGGGGACCC AACACAACAT CCTGAAGATT ATTTAAAGGA AAAGATGCTG 1680
148 ATTGGTACAT AAATCTTTTA CATGGCCTTG GTCTAGAGGA GGCAGGCTTT TAGAATCATG 1740
149 TTTTGTTAAT CCGCTTCACT AAATTGGACC TTCACATATC TAAAAAGCTC TGAAGTGTTC 1800
150 GTATATTTGA AATACCTCAA TAAAGAGAGA GCTCATTGAC TGT 1843

```

(2) INFORMATION FOR SEQ ID NO:3:

RAW SEQUENCE LISTING PATENT APPLICATION US/09/007,306

DATE: 06/09/98
TIME: 08:38:57

INPUT SET: S3103.raw

153

154

(i) SEQUENCE CHARACTERISTICS:

155

(A) LENGTH: 346 amino acids

156

(B) TYPE: amino acid

157

(C) STRANDEDNESS: single

158

(D) TOPOLOGY: linear

159

160

(vii) IMMEDIATE SOURCE:

161

(A) LIBRARY: GenBank

162

(B) CLONE: 473772

163

164

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

165

166	Met	Arg	Ile	Glu	Glu	Asp	Leu	Lys	Leu	Gly	Phe	Lys	Asp	Val	Leu	Ile
167	1			5						10					15	
168	Arg	Pro	Lys	Arg	Ser	Thr	Leu	Lys	Ser	Arg	Ser	Asp	Val	Glu	Leu	Glu
169			20					25					30			
170	Arg	Gln	Phe	Thr	Phe	Lys	His	Ser	Gly	Gln	Ser	Trp	Ser	Gly	Val	Pro
171			35				40						45			
172	Ile	Ile	Ala	Ala	Asn	Met	Asp	Thr	Val	Gly	Thr	Phe	Ser	Met	Ala	Ser
173		50				55					60					
174	Ala	Leu	Ala	Ser	Phe	Asp	Ile	Leu	Thr	Ala	Val	His	Lys	His	Tyr	Ser
175	65				70					75					80	
176	Val	Glu	Glu	Trp	Gln	Ala	Phe	Ile	Asn	Asn	Ser	Ser	Ala	Asp	Val	Leu
177				85					90					95		
178	Lys	His	Val	Met	Val	Ser	Thr	Gly	Thr	Ser	Asp	Ala	Asp	Phe	Glu	Lys
179			100					105					110			
180	Thr	Lys	Gln	Ile	Leu	Asp	Leu	Asn	Pro	Ala	Leu	Asn	Phe	Val	Cys	Ile
181			115				120					125				
182	Asp	Val	Ala	Asn	Gly	Tyr	Ser	Glu	His	Phe	Val	Gln	Phe	Val	Ala	Lys
183		130				135					140					
184	Ala	Arg	Glu	Ala	Trp	Pro	Thr	Lys	Thr	Ile	Cys	Ala	Gly	Asn	Val	Val
185	145				150					155					160	
186	Thr	Gly	Glu	Met	Cys	Glu	Glu	Leu	Ile	Leu	Ser	Gly	Ala	Asp	Ile	Val
187				165				170						175		
188	Lys	Val	Gly	Ile	Gly	Pro	Gly	Ser	Val	Cys	Thr	Thr	Arg	Val	Lys	Thr
189			180					185					190			
190	Gly	Val	Gly	Tyr	Pro	Gln	Leu	Ser	Ala	Val	Ile	Glu	Cys	Ala	Asp	Ala
191			195				200					205				
192	Ala	His	Gly	Leu	Gly	Gly	Met	Ile	Val	Ser	Asp	Gly	Gly	Cys	Thr	Thr
193		210				215						220				
194	Pro	Gly	Asp	Val	Ala	Lys	Ala	Phe	Ala	Arg	Ala	Asp	Phe	Val	Met	Leu
195	225				230					235					240	
196	Gly	Gly	Met	Leu	Ala	Gly	His	Glu	Glu	Ser	Gly	Gly	Arg	Ile	Val	Glu
197				245				250						255		
198	Glu	Asn	Gly	Glu	Lys	Phe	Met	Leu	Phe	Tyr	Gly	Met	Ser	Ser	Glu	Ser
199			260					265					270			
200	Ala	Met	Lys	Arg	His	Val	Gly	Gly	Val	Ala	Glu	Tyr	Arg	Ala	Ala	Glu
201			275				280						285			
202	Gly	Lys	Thr	Val	Lys	Leu	Pro	Leu	Arg	Gly	Pro	Val	Glu	Asn	Thr	Ala
203			290			295						300				
204	Arg	Asp	Ile	Leu	Gly	Gly	Leu	Arg	Ser	Ala	Cys	Thr	Tyr	Val	Gly	Ala
205	305					310					315					320

RAW SEQUENCE LISTING PATENT APPLICATION US/09/007,306

DATE: 06/09/98
TIME: 08:38:59

INPUT SET: S3103.raw

206 Ser Arg Leu Lys Glu Leu Thr Lys Arg Thr Thr Phe Ile Arg Val Gln
207 325 330 335
208 Glu Gln Glu Asn Arg Ile Phe Asn Asn Leu
209 340 345
210

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 345 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 544455

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

224
225 Met Pro Arg Ile Asp Ala Asp Leu Lys Leu Asp Phe Lys Asp Val Leu
226 1 5 10 15
227 Leu Arg Pro Lys Arg Ser Ser Leu Lys Ser Arg Ala Glu Val Asp Leu
228 20 25 30
229 Glu Arg Thr Phe Thr Phe Arg Asn Ser Lys Gln Thr Tyr Ser Gly Ile
230 35 40 45
231 Pro Ile Ile Val Ala Asn Met Asp Thr Val Gly Thr Phe Glu Met Ala
232 50 55 60
233 Ala Val Met Ser Gln His Ser Met Phe Thr Ala Ile His Lys His Tyr
234 65 70 75 80
235 Ser Leu Asp Asp Trp Lys Leu Phe Ala Thr Asn His Pro Glu Cys Leu
236 85 90 95
237 Gln Asn Val Ala Val Ser Ser Gly Ser Gly Gln Asn Asp Leu Glu Lys
238 100 105 110
239 Met Thr Ser Ile Leu Glu Ala Val Pro Gln Val Lys Phe Ile Cys Leu
240 115 120 125
241 Asp Val Ala Asn Gly Tyr Ser Glu His Phe Val Glu Phe Val Lys Leu
242 130 135 140
243 Val Arg Ala Lys Phe Pro Glu His Thr Ile Met Ala Gly Asn Val Val
244 145 150 155 160
245 Thr Gly Glu Met Val Glu Glu Leu Ile Leu Ser Gly Ala Asp Ile Ile
246 165 170 175
247 Lys Val Gly Val Gly Pro Gly Ser Val Cys Thr Thr Arg Thr Lys Thr
248 180 185 190
249 Gly Val Gly Tyr Pro Gln Leu Ser Ala Val Ile Glu Cys Ala Asp Ser
250 195 200 205
251 Ala His Gly Leu Lys Gly His Ile Ile Ser Asp Gly Gly Cys Thr Cys
252 210 215 220
253 Pro Gly Asp Val Ala Lys Ala Phe Gly Ala Gly Ala Asp Phe Val Met
254 225 230 235 240
255 Leu Gly Gly Met Phe Ser Gly His Thr Glu Cys Ala Gly Glu Val Phe
256 245 250 255
257 Glu Arg Asn Gly Arg Lys Leu Lys Leu Phe Tyr Gly Met Ser Ser Asp
258 260 265 270

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/007,306

DATE: 06/09/98
TIME: 08:39:01

INPUT SET: S3103.raw

Line

Error

Original Text